In the Specification

Please add the following paragraph on page 1, line 2, after the title:

-- **CROSS REFERENCE TO RELATED APPLICATIONS**

This application is the U.S. national phase application under 35 U.S.C. § 371 of International Application No. PCT/NZ2004/000153, filed July 19, 2004, published as WO 2005/007178 on January 27, 2005, and claiming priority to New Zealand Application No. 527075, filed July 18, 2003 and New Zealand Application No. 532382, filed April 20, 2004. --

Please amend the paragraph beginning on page 6, line 7, as follows:

-- Polypeptide sequence identity may also be calculated over the entire length of the overlap between a candidate and subject polynucleotide sequences using global sequence alignment programs. EMBOSS-needle (available at ~~http://www.ebi.ac.uk/emboss/align/~~ the website of the European Bioinformatics Institute) and GAP (Huang, X. (1994) On Global Sequence Alignment. Computer Applications in the Biosciences 10, 227-235) are also suitable global sequence alignment programs for calculating polypeptide sequence identity. --

Please amend the paragraph beginning on page 12, line 4, as follows:

-- Multiple sequence alignments of a group of related sequences can be carried out with CLUSTALW (Thompson, J.D., Higgins, D.G. and Gibson, T.J. (1994) CLUSTALW: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, positions-specific gap penalties and weight matrix choice. Nucleic Acids Research, 22:4673-4680, ~~http://www-igbmc.u-strasbg.fr/BioInfo/ClustalW/Top.html~~ see the website of the European Molecular Biology Laboratory) or T-COFFEE (Cedric Notredame, Desmond G. Higgins, Jaap Heringa, T-Coffee: A novel method for fast and accurate multiple sequence alignment, J. Mol. Biol. (2000) 302: 205-217))or PILEUP, which uses progressive, pairwise alignments. (Feng and Doolittle, 1987, J. Mol. Evol. 25, 351). --

Please amend the paragraph beginning on page 12, line 20, as follows:

-- PROSITE (Bairoch and Bucher, 1994, Nucleic Acids Res. 22, 3583; Hofmann *et al.*, 1999, Nucleic Acids Res. 27, 215) is a method of identifying the functions of uncharacterized proteins translated from genomic or cDNA sequences. The PROSITE database

~~(www.expasy.org/prosite)~~ (accessible on the Swiss Institute of Bioinformatics website) contains biologically significant patterns and profiles and is designed so that it can be used with appropriate computational tools to assign a new sequence to a known family of proteins or to determine which known domain(s) are present in the sequence (Falquet *et al.*, 2002, Nucleic Acids Res. 30, 235). Prosearch is a tool that can search SWISS-PROT and EMBL databases with a given sequence pattern or signature. --